Compatibility of the Dispatch Distributed Ledger and Dispatch Artifact Network with the EU General Data Protection Regulation

The potential for technology to transform our daily lives is a perpetually current issue. In recent years, distributed ledger technology (**DLT**) has increasingly garnered public attention worldwide, for a wide variety of data-driven applications. Among other things, DLT can offer greater security, transparency, and trust in outcomes for the transactions and relationships that distributed ledgers can facilitate and memorialize.

At the same time, the European Union (**EU**) has been independently transitioning to a more robust regulatory framework for privacy under the General Data Protection Regulation (**GDPR**), which aims to update EU data protection laws to address new technologies. The GDPR was adopted in 2016 and took effect on May 25, 2018. Yet, DLT technology has been advancing very rapidly over recent years, and adoption of GDPR occurred before DLT gained its current prominence in the market as a potential technical solution to various marketwide issues. As a result, the GDPR did not necessarily take the technology underpinning DLT into account at the time of adoption.[1] The result is a potential disconnect between the GDPR regulatory framework and DLT, which can be a challenge for companies that create or work with distributed ledgers.

We agree that GDPR presents certain challenges for distributed ledgers. However, while there is room for eventual improvement to GDPR, we believe that these challenges can be successfully addressed based on a reasonable interpretation of GDPR as it was adopted. Furthermore, GDPR compliance is not one-size-fits-all. Adoption of proper technical solutions and policies as part of a distributed ledger solution can greatly facilitate GDPR compliance.

In the present paper, Dispatch Labs has teamed with London-based GDPR experts Z/Yen and Lily Innovation to explain how distributed ledgers can address GDPR challenges, and in particular how the technical and policy features of the Dispatch ledger and its associated Dispatch Artifact Network (**DAN**) are designed to address and minimize these challenges.[2]

[1] See, e.g., Blockchain Bundesverband (the German Blockchain Association), "Blockchain, data protection, and the GDPR", p. 2 (May 25, 2018), ("GDPR was created before Blockchain and is already outdated, since it doesn't account for decentralized technologies."), available at: https://www.bundesblock.de/wp-content/uploads/2018/05/GDPR_Position_Paper_v1.0.pdf ("Blockchain Bundesverband Paper").
[2] See Dispatch Labs, "A shared ledger framework for programmable information" (March 5, 2018) (the "Dispatch White Paper"), https://dispatchlabs.io/wp-content/uploads/2018/03/DispatchWhitepaper_Mar_5_18_v1.55.pdf.

We address four principal GDPR challenges that have been identified for distributed ledgers:

- the **permanent availability** of data on distributed ledgers to all who have access to the ledger;
- the **widespread and ongoing processing of ledger data**;
- the potential difficulty of **identifying data controllers** (*i.e.* persons or entities responsible for data processing under GDPR) **and data processors** (*i.e.* persons or entities that process data on behalf of a data controller) in the context of a distributed ledger; and
- the potential use of distributed ledgers for **automated decision-making**.

The first two issues are particular challenges for distributed ledgers and relate directly to the use of an immutable, distributed ledger. The latter two issues apply in a variety of data protection contexts but present some specific issues for distributed ledgers. We consider each of the four issues in turn below.

1. Data Permanence

The most frequently identified tension between GDPR and distributed ledger technology involves the inherent immutability of distributed ledgers. Immutability is a key feature of distributed ledgers, enhancing their security and allowing them to provide a permanent, verifiable, public record of transactions. However, this immutability presents a potential for conflict with GDPR principles, especially:

- the **storage limitation principle** under GDPR Art 5(1)(e) that "personal data shall be … kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed"; and

- the **right to erasure** (also known as the 'right to be forgotten') under GDPR Art 17, which provides an obligation on data controllers to erase data under specified circumstances.

As noted at the beginning of this paper, the main reason for the tension between these principles and distributed ledger immutability appears to be, simply, that the authors of GDPR did not anticipate DLT at the time that GDPR was adopted. But this does not mean that GDPR prevents or seriously impedes the deployment of these technologies, for two main reasons.

First, neither of the above principles is absolute. The storage limitation principle only restricts ongoing storage for "longer than is necessary for the purposes for which the personal data are processed". Since immutability is a fundamental function and feature of DLT, it follows that with adequate advance notice of these functions users of the technology can be considered to have accepted that use inherently involves permanent storage and that this is "necessary for the purposes for which the personal data are processed".

Likewise, the right to erasure applies only in specified circumstances – most importantly where (a) the "personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed", (b) the data subject has withdrawn consent (if processing is based upon consent) or (c) the data controller processes personal data based on its "legitimate interests" without adequate justification.[3] Where permanent storage is *required*, as in the case of distributed ledgers, there are likely to be relatively few circumstances in which these bases for erasure would be applicable. Among other things, processing in the DLT context is frequently justified as necessary to perform a contract with the user[4] (in which case withdrawal of consent is likely to be irrelevant).

Second, it is possible to design a distributed ledger to allow for the effective erasure of personal data, by storing that data off the ledger itself, or alternatively on the ledger in encrypted form that can

only be decrypted by the holder of a private key. The crucial innovation of the Dispatch ledger to facilitate this approach is the separation of data storage onto the DAN. In addition, the Dispatch

[3] GDPR Art. 17(1)(a)-(c). The other circumstances giving rise to a right of erasure involve unlawful processing, legal requirements for deletion and data regarding children. GDPR Art. 17(1)(d)-(f).
[4] GDPR Art. 6(1)(b).

ledger supports Dapps – flexible applications that build and run operations to access, distribute and manipulate data on the ledger.

The Dispatch protocol allows erasure of data by enabling individual Dispatch ledger users[5] (and/or Dapps deployed by users) to choose to render data on the DAN permanently inaccessible by either of the following methods:

- deleting references to the DAN data from the Kademlia distributed hash table (**DHT**) maintained by the 'farmer' node(s) responsible for the data (further details on this approach are available in the Dispatch White Paper, available on the Dispatch Labs website); and/or

- deleting private keys for storing the data – the Dispatch protocols allow users to assign private keys to individual data items (or groups of data items) on the DAN (known as 'artifacts').

Users can deploy these methods flexibly, according to the data protection requirements of their applications, using the code and tools of the Dispatch protocol.

Either of the above methods has essentially the same effect as deleting the data, and together even more so. Indeed, some data protection authorities have already concluded that irreversible encryption constitutes erasure[6], and both of these methods involve irreversible encryption – since data on the DAN is encrypted and cannot be decrypted when either the pointer to it (in the Kademlia DHT) or the private key is not available.

[5] We use 'users' generally to refer to any entity or individual building an application on the Dispatch ledger and the DAN ('application developers'), as well as entities and individuals using such applications ('end users'). In practice for a particular application, decisions to render data inaccessible may be made by applications developers or end users or both, depending upon the specific implementation chosen by the application developer.

[6] *See* Hogan Lovells, "A guide to blockchain and data protection", p. 15 (September 2017), https://www.hlengage.com/_uploads/downloads/5425GuidetoblockchainV9FORWEB.pdf.

2. Widespread and Ongoing Processing of Ledger

Although distributed ledgers have been used for decades[7], their recent explosive growth was initially driven by the Bitcoin protocol, which allows trust-free transaction verification through a proof of work consensus protocol.[8] This protocol requires every node of the Bitcoin network that wishes to engage in 'mining' of new bitcoins to repeatedly process new blocks of transactions (each time with a different 'nonce', or padding data) using a hash algorithm, until the calculated hash is below a certain value. Some have argued that this repeated processing – to the extent the processed blocks include personal data – conflict with GDPR principles, especially:

- the **data minimisation principle** under GDPR Art 5(1)(c) that "personal data shall be … limited to what is necessary in relation to the purposes for which they are processed"; and

- the **right to restriction of processing** under GDPR Art. 18 and the **right to object to processing** under GDPR Art. 21, which require data controllers to terminate or restrict the processing of personal data upon request in certain circumstances.

The conflict between these provisions and distributed ledgers is less obvious than with respect to data permanence. Furthermore, similar to the case of data permanence, there are two main bases for addressing any legal concerns.

First, the principles themselves are subject to significant limitations. With respect to the data minimisation principle, there is a good argument that multiple hashing of personal data does not mean that such data are not "limited to what is necessary in relation to the purposes for which they are processed". That is, multiple hashing does not increase the *amount of personal data* that is processed – which is the core focus of the data minimisation principle – but rather relates to the *number of times* that the data are processed.

Likewise, the right to restriction of processing and right to object to processing apply only in specified circumstances, which are narrower than those triggering the right to erasure, i.e. where (a) there is a challenge to processing based upon "legitimate interests" (as for the right to erasure), (b) there is a challenge to accuracy of personal data, (c) processing is unlawful, (d) the data controller no longer needs the data but the data subject (*i.e.* the individual to whom the data relate) wishes the data to be retained for reasons related to legal claims, or (e) the processing involves use of profiling for direct market.[9] On a distributed ledger, there may be no way to entirely stop processing of the ledger in any of these circumstances; however, it is entirely possible to design distributed ledger solutions so that any personal data is encrypted and cannot be processed in a manner that discloses the data in these circumstances. The Dispatch ledger is designed to accommodate any computer application, and this allows implementation of such limitations. Examples of such limitations include use of the DAN to limit processing via off-ledger storage (as explained in section 1 above)

[7] See Arvind Narayanan & Jeremy Clark, "Bitcoin's Academic Pedigree", ACM Queue (Aug. 29, 2017), https://queue.acm.org/detail.cfm?id=3136559 ("Bitcoin's ledger data structure is borrowed, with minimal modifications, from a series of papers by Stuart Haber and Scott Stornetta written between 1990 and 1997 (their 1991 paper had another co-author, Dave Bayer).").
[8] *See* Bitcoin Wiki, "Proof of work", https://en.bitcoin.it/wiki/Proof_of_work.
[9] GDPR Arts. 18(1), 21(1) & 21(2).

and enabling certain users to operate "light" nodes that act as a wallet and process only the transactions for a particular user (and not other ledger data).

Second, not all distributed ledger protocols are created equal. The Dispatch ledger uses the delegated asynchronous proof-of-stake (**DAPoS**) consensus protocol. As explained in the Dispatch White Paper, DAPoS involves significantly less frequent processing of ledger transactions than a proof-of-work consensus protocol. Use of DAPoS materially reduces any unnecessary processing that could be restricted by the data minimisation principle. Furthermore, as explained in the previous paragraph, the flexible programming architecture of the Dispatch ledger and Dapps allows limits on processing of personal data, in order to implement the right to restriction of processing and right to object to processing.

3. Identifying Data Controllers and Data Processors

Unlike the previous two issues, the challenge of identifying data controllers and data processors is not specific to distributed ledgers. GDPR defines 'controller' as:

> *"the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data".*[10]

'Processor' is defined as:

> *"a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller"*

Applying these definitions to complex, multi-party applications and technical ecosystems (consider, for example, the interactions of buyers, sellers and payment providers on a platform like Amazon or eBay) is a frequent challenge for data protection practitioners. However, despite potential ambiguities, it is our experience that a practical, good faith approach to defining controller and processor roles is generally effective from a regulatory perspective.

Dispatch Labs has taken a straightforward approach to defining who is a data controller and who is a data processor in the Dispatch ecosystem:

- users that store data and build applications on the Dispatch ledger are data controllers with respect to any personal data that they process or store;
- Dispatch nodes (see the Dispatch White Paper for information on the different types of nodes) are data processors – this includes Dispatch Labs when it acts as a node; and
- Dispatch Labs is a data controller with respect to personal information on customers, users, partners, developers and the general Dispatch community that it collects for other business purposes (see the Dispatch Labs Privacy Policy for further information).

This approach to defining roles of data controllers and data processors is consistent with approaches recently recommended by the German Blockchain Association[11] and others.

[10] GDPR Art 4(7).
[11] *See* Blockchain Bundesverband Paper, pp. 5-7.

4. Automated Decision-Making

Like the previous issue, the question of automated decision-making is not specific to distributed ledgers. Article 22 of GDPR restricts automated decision-making without human involvement "which produces legal effects concerning [an individual] or similarly significantly affects him or her." This provision has generated substantial interest and concern in the technology community because a wide variety of emerging applications – particularly those involving artificial intelligence and machine learning – use automated decision-making.[12]

For distributed ledgers, the Article 22 restriction is only relevant to a distributed ledger application to the extent that the application uses automated decision-making – whether any application in fact does so must be assessed on a ledger-specific and application-specific basis.

At the protocol / architecture level, the Dispatch ledger and the DAN are designed to avoid any significant issue under Article 22. Under the DAPoS protocol, voting for delegates who validate transactions involves human decisions, and so is not implicated by the Article 22 limitations. Automated decision-making is used only for:

- actual approval of transactions – this simply validates voluntary actions taken by the transacting parties, so does not appear to produce material "legal effects … or similar[] significant[] [e]ffects" and in many cases will be within the exception of GDPR Art. 22(2)(a) for processing that "is necessary for entering into, or performance of, a contract between the data subject and a data controller"; and

- security (*i.e.* monitoring node behaviour and taking action in response to anomalous behaviour) – this is to protect users, and is essentially unrelated to legal or similar effects.

Of course, applications built on the Dispatch ledger can use automated decision-making, and such decision-making may be subject to Article 22. This can be the case for any computer application, whether or not built on a distributed ledger. The flexible programming environment provided by the Dispatch ledger and DAN is designed to allow users to adopt effective compliance strategies in such cases.

[12] *See, e.g.*, Pomin Wu, "GDPR and its impacts on machine learning applications", Medium (Nov. 7, 2017), https://medium.com/trustableai/gdpr-and-its-impacts-on-machine-learning-applications-d5b5b0c3a815.

Conclusions

The analysis in this paper demonstrates both that GDPR and DLT are not incompatible, and that the technology features of the Dispatch ledger and the DAN make them especially well-suited to supporting GDPR-compliant DLT solutions.

In the current Internet environment where use of personal data is a major driver of business models, consumers and governments are increasingly concerned about misuse of personal data. GDPR is a major step to address these concerns in Europe, and there are signs that it will become a privacy 'gold standard' that will also be adopted elsewhere. For example, the new privacy law passed in California at the end of June 2018 draws much from GDPR.

Although these changes to privacy law can present challenges for technology companies, there is no question that the technology sector can and will find a way through the challenges of GDPR and other privacy law. There is no sign that GDPR is materially impeding innovation and growth in the sector generally, or for DLT and blockchain solutions. Although there are challenges under GDPR for DLT solutions, there are also clear ways through those challenges, which we have explored in this paper.

Dispatch Labs has chosen to make a robust response to such privacy challenges into a core feature of its business model. Specifically, Dispatch Labs is taking GDPR and other privacy issues into account in developing its technologies, including through the research and internal discussion associated with this paper.

Given the careful technology choices made by Dispatch Labs, the Dispatch ledger is especially well-suited for addressing GDPR compliance issues. As a result, the Dispatch ledger, the DAN and Dapps built on them can support data solutions providing privacy guarantees at least equal and often stronger than those offered by other data infrastructure technologies.